

## RESEARCH ARTICLE

# Association between *IL1* gene polymorphism and human African trypanosomiasis in populations of sleeping sickness foci of southern Cameroon

Elvis Ofon<sup>1</sup>, Harry Noyes<sup>2</sup>, Vincent Ebo'o Eyanga<sup>3</sup>, Flobert Njiokou<sup>4</sup>, Mathurin Koffi<sup>5</sup>, Pythagore Fogue<sup>1</sup>, Christiane Hertz-Fowler<sup>2</sup>, Annette MacLeod<sup>6</sup>, Enock Matovu<sup>7</sup>, Gustave Simo<sup>1\*</sup>, for the TrypanoGEN Research Group, as members of The H3Africa Consortium<sup>†</sup>



**1** Molecular Parasitology and Entomology Unit, Department of Biochemistry, Faculty of Science, University of Dschang, Dschang, Cameroon, **2** Centre for Genomic Research, University of Liverpool, Liverpool, United Kingdom, **3** MINSANTE, Divisional Centre for Diseases, PNLTHA, Ministry of Public Health, Yaoundé, Cameroon, **4** Laboratory of Molecular Biology, Department of Animal Biology, Faculty of Science, University of Yaoundé I, Yaoundé, Cameroon, **5** Université Jean Lorougnon Guédé (UJLoG), UFR Environnement-Santé, Laboratoire des Interactions Hôte-Microorganismes-Environnement et Evolution (LIHME) Daloa, Côte d'Ivoire, **6** Wellcome Centre for Molecular Parasitology, University Place, Glasgow, United Kingdom, **7** College of Veterinary Medicine, Animal Resources and Bio-security, Makerere University, Kampala, Uganda

 OPEN ACCESS

**Citation:** Ofon E, Noyes H, Ebo'o Eyanga V, Njiokou F, Koffi M, Fogue P, et al. (2019) Association between *IL1* gene polymorphism and human African trypanosomiasis in populations of sleeping sickness foci of southern Cameroon. *PLoS Negl Trop Dis* 13(3): e0007283. <https://doi.org/10.1371/journal.pntd.0007283>

**Editor:** Guy Caljon, Universiteit Antwerpen, BELGIUM

**Received:** October 23, 2018

**Accepted:** March 7, 2019

**Published:** March 25, 2019

**Copyright:** © 2019 Ofon et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its Supporting Information files.

**Funding:** We are grateful to the University of Dschang as well as the Wellcome Trust for funding this work (099310/Z/12/Z) as part of its support for the H3Africa initiative (<http://h3africa.org/>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

<sup>†</sup> Membership of the TrypanoGEN Research Group, as members of the H3Africa Consortium is provided in the Acknowledgments.

\* [gsimoca@yahoo.fr](mailto:gsimoca@yahoo.fr), [gustave.simo@univ-dschang.org](mailto:gustave.simo@univ-dschang.org)

## Abstract

### Background

Human African Trypanosomiasis (HAT) is a neglected tropical disease caused by infections due to *Trypanosoma brucei* subspecies. In addition to the well-established environmental and behavioural risks of becoming infected, there is evidence for a genetic component to the response to trypanosome infection. We undertook a candidate gene case-control study to investigate genetic associations further.

### Methodology

We genotyped one polymorphism in each of seven genes (*IL1A*, *IL1RN*, *IL4RN*, *IL6*, *HP*, *HPR*, and *HLA-G*) in 73 cases and 250 controls collected from 19 ethno-linguistic subgroups stratified into three major ethno-linguistic groups, 2 pooled ethno-linguistic groups and 11 ethno-linguistic subgroups from three Cameroonian HAT foci. The seven polymorphic loci tested consisted of three SNPs, three variable numbers of tandem repeat (VNTR) and one INDEL.

### Results

We found that the genotype (TT) and minor allele (T) of *IL1A* gene as well as the genotype 1A3A of *IL1RN* were associated with an increased risk of getting *Trypanosoma brucei*

**Competing interests:** The authors have declared that no competing interests exist.

*gambiense* and develop HAT when all data were analysed together and also when stratified by the three major ethno-linguistic groups, 2 pooled ethno-linguistic subgroups and 11 ethno-linguistic subgroups.

## Conclusion

This study revealed that one SNP rs1800794 of *IL1A* and one VNTR rs2234663 of *IL1RN* were associated with the increased risk to be infected by *Trypanosoma brucei gambiense* and develop sleeping sickness in southern Cameroon. The minor allele T and the genotype TT of SNP rs1800794 in *IL1A* as well as the genotype 1A3A of *IL1RN* rs2234663 VNTR seem to increase the risk of getting *Trypanosoma brucei gambiense* infections and develop sleeping sickness in southern Cameroon.

## Author summary

Human African Trypanosomiasis (HAT), or sleeping sickness, is a parasitic disease caused by flagellated parasites of the genus *Trypanosoma*. This disease has been included into the WHO roadmap for neglected tropical diseases with elimination as a public health problem targeted for 2020 and the interruption of transmission to humans for 2030. To achieve these elimination and interruption goals, it is important to identify and understand the factors that may hamper these goals. Understanding the contribution of human genetics to the response of trypanosome infections is important for the development of new control strategies. In this study, polymorphism in seven genes was investigated between controls and sleeping sickness patients of three sleeping sickness foci of Southern Cameroon in order to see if there is any association with the development of disease. Results of this study have shown that the genotype (TT) and minor allele (T) of *IL1A* gene and the genotype 1A3A VNTR of *IL1RN* are associated with an increased risk of getting *T. b. gambiense* infections and develop sleeping sickness in major ethno-linguistic groups of the Cameroonian population. They suggest that the association between host genetic determinants and the susceptibility to *T. b. gambiense* infections could vary according to the population studied. These results will improve our knowledge on the role of human genetics determinants and the risk to be infected by *T. b. gambiense* and develop sleeping sickness. They could thus lead to the identification of novel biomarkers which could open a frame work for the development of new diagnostics, treatments and intervention strategies.

## Introduction

Human African Trypanosomiasis (HAT), or sleeping sickness, is a parasitic infection caused by flagellated parasites of the genus *Trypanosoma*. The parasites belong to *Trypanosoma brucei* complex which is subdivided into three subspecies: *Trypanosoma brucei gambiense* (Dutton, 1902) is responsible for the chronic form of the disease in West and Central Africa, *T. b. rhodesiense* (Stephen and Fantham, 1910) causes the acute form of HAT in East and South Africa whilst *T. b. brucei* is only infective to animals. These trypanosomes are transmitted through the bites of haematophagous flies of the genus *Glossina* commonly known as tsetse flies [1, 2]. HAT was considered to be under control during the 1960s, but the disease has re-emerged in the last decades as a public health problem in many sub-Saharan African countries due to the

abandonment of control measures after independence and also to socio-political and environmental upheavals [2]. HAT is often fatal unless treated and is endemic in 36 sub-Saharan Africa countries with about 65 million people in more than 250 foci exposed to the risks of infections [3]. Currently, more than 98% of the reported cases are due to *T. b. gambiense* infection [3]. Control efforts undertaken by the national sleeping sickness control programs have succeeded in considerably reducing the number of cases, and less than 2200 new cases were officially reported in 2016 [4]. Sleeping sickness has been included in the WHO roadmap for neglected tropical diseases with elimination as a public health problem targeted for 2020, and the interruption of transmission to humans for 2030.

To achieve the elimination and interruption goals, it is important to identify and gain a better understanding of the clinical evolution of the disease and the factors that may hamper this goal. Addressing the contribution of human genetics to the response to *T. b. gambiense* infections is important for the development of new control strategies because a range of clinical presentations of HAT including asymptomatic carriers and spontaneous cure without treatment have been reported in West Africa [5]. Understanding the genetic bases of these new disease profiles may help to identify more susceptible populations for more effective control operation. Previous studies have identified polymorphisms in *APOL1*, *IL6*, *HLA-G* and *HP/HPR* that regulate the human susceptibility to trypanosome infections [6, 7, 8, 9, 10, 11, 12, 13]. Most of these genes seem to play important roles during *T. b. gambiense* infection [14]. For instance, *HLA-G* has been reported to be involved in HAT progression. In addition, *IL1* participates in macrophage activation during early *T. b. gambiense* infection in mice [15]. It also plays a key role in the recruitment of leukocytes into the CNS during *T. b. gambiense* infections [16, 17]. However, contrasting results on the association between gene polymorphisms and the risk to be infected by *T. b. gambiense* and develop HAT have been reported between countries and therefore, efforts are needed to better understand the genetic bases of human susceptibility to *T. b. gambiense* infections. In this current study, polymorphisms in seven genes were genotyped to identify any association with HAT. Our data suggest that the association between host genetic determinants and the susceptibility to be infected by *T. b. gambiense* and develop sleeping sickness could vary according to the population studied.

## Materials and methods

### Study areas

This study was conducted in three active sleeping sickness foci in the forest region of Southern Cameroon. The three HAT foci were Bipindi and Campo in the Southern region and Fontem in the South-west region of Cameroon.

The Campo focus (2° 82'00"N, 9° 85'20"E) is located in the equatorial forest and extends from the Atlantic coast along the Ntem river which delimits the Cameroon–Equatorial Guinea border. It is a hypo-endemic focus where no epidemic outbreak has been observed for many decades [18]. It is a cosmopolitan area with several ethnic groups (mainly the Iyassa, Kwasse, Maabi, Mvae and Ngoumba) with most of them speaking Bantu family languages. Other minor ethnic groups are semi Bantus, Sao-Sudanese and Baka [19].

The Bipindi HAT focus (3° 82'00"N, 10° 82'20"E) is located at about 75 km from the Atlantic coast in the South of Cameroon. Bipindi has been known as a HAT focus since 1920. During the last two decades, it was among the most active HAT foci of Cameroon with about 83 HAT cases diagnosed from 1999 to 2011 [19]. About 95% of the inhabitants of the Bipindi HAT focus are Bantu speakers and the majority belong to the Ngoumba, Nti and Fan. The remaining 5% of inhabitants are Baka, semi Bantus and Sao-Sudanese speakers.

The Fontem focus (5°40'00"N, 9°55'00"E) is located in the South-West Region of Cameroon where HAT has been known to occur since 1949 [20]. It was previously among the most active HAT foci of Cameroon [21], but in recent decades, it has become hypo-endemic with about 8 patients detected among 16,000 persons examined between 1998 and 2007 [22]. In this focus, the Mundani and Banyangi are the major ethnic groups. Other minor ethnic groups such as Bangwa and Bamileke are also found.

### Study population

The Cameroonian population is made up of more than 250 ethno-linguistic subgroups from three major ethnic groups: Bantu (e.g.: Bulu, Bassa, Bakundu, Maka, Douala), Semi Bantu (e.g.: Bamileke, Gbaya, Bamoun, Tikar) and Sudano-Sao (e.g.: Fulbe, Mafa, Toupouri, Shoa-Arabs, Moundang, Massa, Mousgoum) [11] (S1 Table). Beside these three groups, some minor groups exist such as the Baka who generally speak the Bantu languages but who are not closely related to any of these three major groups [23]. The ethnic composition varies considerably between regions, HAT foci and even within the same HAT focus.

### Ethical consideration

The protocol of this study was approved by the Ethical Committee of the Ministry of Public Health of Cameroon on 21 November 2013 with a reference number N°2013/11/364/L/CNERSH/SP. The local administrative and traditional authorities of each HAT focus were also informed and gave their approval. Subsequently, the review board of the Laboratory of Microbiology and Anti-microbial Substances (LAMAS) of the Department of Biochemistry of the Faculty of Science of the University of Dschang gave its approval. All adult subjects provided informed consent, and a parent or guardian of any child participant below 18 years old provided informed consent on their behalf. Each informed consent was written because all individuals enrolled in this study gave their approval by signing an informed consent form and a Certificate of Confidentiality. In addition, an assent form was also signed by children below 18 years old. During analyses, data for each subject were anonymized.

### Sample collection

Blood samples were collected during medical surveys performed jointly with the National Sleeping Sickness Control Program of Cameroon and the research team of the molecular parasitology and entomology unit of the University of Dschang. The sampling was performed in Campo in 2014 and 2017, in Bipindi in 2015 and 2017, and in Fontem in 2015. During these surveys, all participants at risk were tested using the Card Agglutination Test for Trypanosomiasis (CATT). It was performed on blood collected by finger prick [24]. This immunological test was carried out to screen people who have been in contact with *T. b. gambiense*. It was initially performed on whole blood as described by Magnus *et al.* [24]. For all participants with a positive CATT on whole blood, blood sample was collected in EDTA tubes and a two-fold plasma dilution series in CATT buffer was tested to assess the end titer, i.e. the highest dilution still positive on plasma (CATT-P). All individuals with CATT dilution on plasma  $\geq 1/8$  underwent parasitological examinations by direct examination using the capillary tube centrifugation (CTC) [25] and mini-anion exchange centrifugation technique (mAECT) [26]. Beside the CTC and mAECT, lymph node aspiration followed by a microscopic examination was performed to search for trypanosomes in all individuals showing enlarged lymph nodes. All controls and participants with a CATT dilution  $\geq 1/8$  and negative for parasitological tests were subjected to the trypanolysis test in order to confirm their status [27].

## Trypanolysis test

Ninety micro-liters of plasma sample from each control and each individual with CATT dilution  $\geq 1/8$  and negative for parasitological tests were spotted on a Whatman paper disc (divided in three equal parts with each bearing a spot of 30 $\mu$ l) that was sent to the “Centre International de Recherche-Développement sur l’Élevage en Zones Sub-humides (CIRDES, Bobo-Dioulasso, Burkina Faso)”. Each plasma sample was tested by the immune trypanolysis test as described by Jamonneau *et al.* [27]. It is the highly specific test for *T. b. gambiense* and constitutes a routine test for the surveillance of HAT. This test was performed on plasma as previously described by Van Meirvenne *et al.* [28] with LiTat 1.3, 1.5, and 1.6 variable antigen types (VAT).

During medical surveys, each new HAT case was defined as an individual in whom trypanosomes were seen by at least one parasitological method. Old HAT cases were also sampled. They were residents in whom trypanosomes had been previously seen by at least one parasitological test after passive or active case detection. Old HAT cases were included only if the information regarding the clinical status, the CATT and all parasitological tests were available in hospital records and in the National control program register.

Each HAT case was matched to at least three controls. This matching was done by age, sex, occupation and when possible by ethno-linguistic subgroup. A control was considered as any individual who was negative for the CATT and trypanolysis tests and all parasitological tests including CTC, mAECT and lymph node examination [11]. These controls were enrolled during medical surveys.

## DNA extraction

Five millilitres of blood were centrifuged at 3500g for 3 minutes and the buffy coat was collected. Genomic DNA was extracted from the Buffy-coat with the QIAamp DNA Blood Midi/Maxi kit (Qiagen) according to the manufacturer’s instructions. The DNA was eluted with 200 $\mu$ l of sterile water and stored at -20°C until use.

## Selection of candidate genes, SNPs, VNTR and INDEL

For this study, seven genes containing three SNPs, three VNTRs and one INDEL (Table 1) were identified and selected based on literature searches. The selected genes and loci were associated to HAT and other diseases. The selection of *HLA-G*, *HP* and *HPR* genes as well as two different cytokines genes (*IL6* and *IL1A*) was based on their previously reported association with HAT [6, 7, 9, 10]. *HPR* and haptoglobin (*HP*) are involved in the lysis of trypanosomes and the scavenging of haem during trypanosome infections [8]. The *IL1* gene has been shown to enhance immune-modulating and stimulating effects on the TLF components and inflammatory immune response activities during HAT infection [29, 30, 31]. Associations between some polymorphic variants within these genes and the outcome of HAT have also been previously outlined [6, 32, 33]. Loci on these genes were selected after literature searches as follows: the SNPs rs1800794 of *IL1A* and rs1554606 of *IL6* and the INDEL rs371194629 in the 3’ UTR of *HLA-G* gene were selected due to their previously reported association with HAT in the DRC although *IL1A* was not associated with HAT [6, 7]. The SNP rs1679370 of *HPR* gene was selected from a study on CNV of associated with HAT [9]. Other genes such as *IL1RN*, *IL4R* and *HP* were also selected not only for their association with other diseases, but especially because HAT seems to trigger inflammatory and immunological responses with biological pathways associating the selected genes [34, 35, 36]. The polymorphic locus rs2234663 within *IL1RN* was selected for its association with *H. Pylori* gastric infections in Brazil [36] while rs79071878-*IL1RN* within *IL4R* gene was due to its association with type II diabetes in

**Table 1. Candidate genes and SNPs, VNTR and indel loci identified and selected for this study.**

Gene	SNP	VNTR	INDEL	Localization (GRCh37)	Characteristics	Reference
<i>IL1A</i>	rs1800794	-	-	chr2:113543273	mRNA Upstream gene variant 2kb	[6]
<i>IL6</i>	rs1554606	-	-	chr7:22768707	Intron variant	[6]
<i>HPR</i>	rs1697370	-	-	chr16:35339932	Intron variant	[9]
<i>IL1RN</i>	-	rs2234663	-	chr2:113888106	Intron variant	[36]
<i>IL4RN</i>	-	rs79071878	-	chr5:132680584	Intron variant	[37]
<i>HP</i>	-	HP1/2 alleles	-	-	VNTR	[38]
<i>HLA-G</i>	-	-	rs371194629	chr6:29798581	3' UTR variant	[7]

SNP: Single Nucleotide Polymorphism; VNTR: Variable Number Tandem Repeats; INDEL: Insertion and Deletion; UTR: Untranslated region; GRCh37: Reference human genome build version.

<https://doi.org/10.1371/journal.pntd.0007283.t001>

India [37]. HP1/2 VNTR allele of the *HP* gene was selected based on its associations with malaria [36].

### Genotyping of SNPs in *IL6*, *IL1A* and *HPR* by PCR-RFLP

In this study, the SNPs in *IL6*, *IL1A* and *HPR* were investigated by PCR-RFLP where a DNA fragment of each of these genes was amplified and subsequently digested by a specific restriction enzyme. The following primer pairs were used: IL6-PF (GTCAAATGTTTAAAACCTCCC ACAGGTT) and IL6-PR (GCAGCCAGAGAGGGAAAAGG) for *IL6* [6], IL1-P-PF (GGCCA CAGGAATTATAAAAGCTGAGA) and IL1-P-PR (GGGAGAAAGGAAGGCATGGATTTT) for *IL1A* [6] and Hpr-F (GAGCCACAAATTCTGACGAG) and Hpr-R (TTGAGGTTCTTGA GGCATT) for *HPR*. The primers Hpr-F and Hpr-R for *HPR* were designed with the Primer3 vs 4.1 software [39, 40].

For each of these three genes, the amplification reactions were performed in a final volume of 25 µl containing 1X of PCR buffer, 1.5 mM MgCl<sub>2</sub>, 20 pmol of each primer, 0.5 units of Taq DNA polymerase (Qiagen) and 5–10 ng of genomic DNA. The amplification program contained a denaturing step at 95°C for 5 min followed by 35 amplification cycles of 95°C for 45 s,

**Table 2. Expected number and size of DNA fragments generated at each locus genotyped.**

Markers	Gene	Locus	Size of DNA fragments in base pair			Reference
			Heterozygote genotype	Homozygote		
				wild type	mutant	
SNPs	<i>IL1A</i>	rs1800794	236/489/725	236/489	725	[6]
	<i>IL6</i>	rs1554606	315/543/858	315/543	858	[6]
	<i>HPR</i>	rs1697370	88/147/235	88/147	235	[9]
VNTR	<i>IL1RN</i>	rs2234663	410/335 410/500 410/240 410/595	410	335 500 240 595	[36]
	<i>IL4RN</i>	rs79071878	183/253	183	253	[37]
	<i>HP</i>	-	1757/3481	1757	3481	[38]
INDEL	<i>HLA-G</i>	rs371194629	345/359	359	345	[7]

SNP: Single Nucleotide Polymorphism; VNTR: Variable Number Tandem Repeats; INDEL: Insertion and Deletion; IL: Interleukin 6; HP: Haptoglobin; HPR: Haptoglobin related protein.

<https://doi.org/10.1371/journal.pntd.0007283.t002>



63°C (*IL6* and *IL1A*) or 60°C (*HPR*) for 60 s and 72°C for 60 s. A final extension step was performed at 72°C for 5 min.

PCR products were visualised by electrophoresis on 2% agarose gel containing ethidium bromide. Ten micro-litters of *IL6*, *IL1A* and *HPR* PCR products were digested with Hind III, Nco I and Bci VI respectively (all enzymes from Thermo Fisher Scientific). The digestion was done overnight at 37°C in the buffer 3.1 provided by the manufacturer. The digested products of *IL6* and *IL1A* were separated by electrophoresis on a 2% agarose gel at 100 volts for 1 h 30 min. For *HPR*, the digested products were resolved by electrophoresis on 3.5% agarose gel at 100 volts for 2 hours.

For rs1554606 of *IL6* and rs1800794 of *IL1A*, three different profiles were expected (Table 2): the homozygote wild type genotype with two DNA fragments of 236 and 489 bp for *IL1A*, 315 and 543 bp for *IL6*, the heterozygote genotype showing three DNA fragments of 236, 489 and 725 bp for *IL1A*, and 315, 543 and 858 bp for *IL6*, and the homozygote genotype with one DNA fragment of 725 bp for *IL1A* and 858 bp for *IL6*.

For rs1697370 of *HPR*, three different profiles were expected (Table 2): the homozygote wild type genotype with two DNA fragments of 88 and 147 bp, the heterozygote genotype showing three DNA fragments of 88, 147 and 235 bp, and the homozygote mutant genotype with one DNA fragment of 235 bp.

To minimize misinterpretation of heterozygote frequency that could result from partial digestion, the amplified product of each sample (control and HAT case) was quantified and the same amount of DNA was subjected to restriction enzyme digestion. Between different amplification and digestion series, an internal control made of sample with known genotype was added. This sample was used to control the reproducibility and digestion efficiency between different amplification and digestion series.

### Amplification of 70 bp and 86 bp tandem repeats of *IL4RN* and *IL1RN* genes

The 70 bp tandem repeat (rs79071878) region of *IL4RN* gene was amplified with IL4-70 bp-F (AGGCTGAAAGGGGAAAGC) and IL4-70 bp-R (CTGTTCACTCAACTGCTCC) primers [37] while the 86bp tandem repeat (rs2234663) of *IL1RN* gene was amplified with IL1RN-F (CTCAGCAACTCCTAT) and IL1RN-R (TCCTGGTCTGCAGGTAA) primers as described by Santos *et al.* [36]. For these two genes, the PCR reactions were performed in a final volume of 25 µl contained 5–10 ng of DNA, 2.5 mM and 2 mM MgCl<sub>2</sub> for *IL4* and *IL1RN* respectively, 0.2 mM of each dNTP, 20 pmol of each primer and 0.5 units of Taq polymerase (Qiagen). The amplification program was 95°C for 5 min followed by 35 cycles of 95°C for 45 s, 61°C for 45 s and 72°C for 60 s. A final extension was performed at 72°C for 5 min. PCR products were separated by electrophoresis on a 2% agarose gel at 100 volts for 1 h 30 min.

The size and number of tandem repeats were evaluated for each sample. For *IL4RN*, the PCR products of 183bp (two repeats of 70b p) and 253 bp (three repeats of 70 bp) correspond to homozygote wild type (genotype R1R1) and homozygote mutant (genotype R2R2) respectively (Table 2). Sample with two DNA fragments of 183 bp and 253 bp was considered as a heterozygote with genotype R1R2.

For *IL1RN*, different alleles with specific sizes could be generated after electrophoresis as described by Santos *et al.* [36]: alleles 1–4 (410 bp), 2–2 (240 bp), 3–5 (500 bp), 4–3 (335 bp) and 5–6 (595 bp). Samples showing one DNA fragment or one allele of 410 bp were considered as homozygote wild type and those with one DNA fragment at 240, 335, 500 or 595 pb were homozygote mutants. Samples presenting two DNA fragments with 410 bp and another one were considered as heterozygote (Table 2).

### Amplification of variable number of tandem repeats of *HP* gene

Genotyping the *HP* polymorphism was performed using two approaches: a PCR approach described by Koch *et al.* [38] and PCR-RFLP using two restriction enzymes to confirm results obtained by PCR [38].

The direct PCR approach consists of two separate PCR reactions with specific DNA fragment characterizing each genotype. Primers A/B (GAGGGGAGCTTGCCTTTCCATTG and GAGATTTTTGAGCCCTGGCTGGT) amplified DNA fragments of 1,757 bp and 3,481 bp for homozygote wild type (genotype Hp1/1) and homozygote mutant (genotype Hp2/2) respectively. Samples showing two DNA fragments at 1,757 bp and 3,481 bp were considered as heterozygote with genotype Hp1/2. Since the 3,481 bp fragment might not amplify due to lower efficiency of PCR for large fragments or sheared genomic DNA, the results were subsequently validated by a second amplification with primers C (CCTGCCTCGTATTAAGTGCACCAT) and D (CCTGCCTCGTATTAAGTGCACCAT), which amplify a specific DNA fragment of 349 bp for the Hp2 allele [38].

For each of these pairs of primers, the PCR reactions were carried out in a final volume of 25 µl containing 5–10 ng of DNA, 2.5 mM MgCl<sub>2</sub>, 0.2 mM of each dNTP, 20 pmol of each primer and 0.5 units of Taq polymerase (Qiagen). The amplification program was 95°C for 5 min followed by 35 cycles of 95°C for 60 s, 69°C for 90 s (primers A/B) or 60 s (primers C/D) and 72°C for 2 min. A final extension was done at 72°C for 5 min. The amplified products were separated by electrophoresis on a 2% agarose gel at 100 volts for 1 h 30 min.

To confirm results (alleles of Hp1 and Hp2) obtained by PCR, the DNA fragments of 1757 bp and 3481 bp of primers A/B were digested with *MspI*, while the fragment of 349 bp of primers C/D was digested with *DraI*. Briefly, 10 µl of amplified DNA fragments of each of the primers set was digested with *MspI* or *DraI* as recommended by the supplier (Thermo Fisher). The digestion was done overnight at 37°C in the buffer 3.1 provided by the manufacturer. The digested products were separated by electrophoresis on a 2% agarose gel at 100 volts for 2 h 30 min.

### Polymorphism in *HLA-G* genes through the analysis of 14bp Indel

The polymorphism at 3'UTR (rs371194629) of *HLA-G* was evaluated by PCR as described by Castelli *et al.* [41]. PCR reactions were performed in a final volume of 25 µl containing 0.2 mM of each dNTP, 1.5 mM MgCl<sub>2</sub>, 20 pmol of each primer (*HLA-G8F*: TGTGAAACAGCTGCC CTGTGT and *HLA-G8R*: GTCTTCCATTTATTTGTCTCT), 0.5 unit of Taq polymerase (Qiagen) and 5–10 ng of genomic DNA. The amplification program was 95°C for 5 min followed by 35 cycles. Each of these cycles was made up of 95°C for 45 s, 56°C for 45 s and 72°C for 1 min. A final extension was performed at 72°C for 5 min.

The amplified products were resolved by electrophoresis on 4% agarose gel for 4 hours at 100 volts. After this resolution, homozygote mutant and homozygote wild type genotypes were identified through DNA fragments of 345 bp and 359 bp for deletion (Del) and insertion (Ins) alleles respectively. For heterozygote genotype, two DNA fragments of 345 bp and 359 bp were expected.

### Statistical analysis

For this study we assumed an additive genetic model where two risk alleles of a SNP (homozygous) have twice the effect of one risk allele (heterozygous) [42]. Power calculation was undertaken using the PGA modeller package in MATLAB software [42]. For this package, the power was calculated by considering an odd ratio (OR) or relative risk (RR) >2 for loci with disease allele frequencies of 0.052–0.500 with 7 loci genotyped. Other factors taken into consideration



include the disease prevalence estimated at  $<0.01$  [43], the standard linkage parameter ( $r^2$ ) for Linkage disequilibrium (LD) of 0.7 [42], a type 1 error of 5% risk and sampling size. This later was estimated as described by Kasiulevicius et al. [44] using the independent case-control sampling size formula [44]. For this estimation, we assumed an expected exposure proportions in control of 0.20, a disease prevalence of  $< 0.01$  [43] and a case-control ratio of 1:3. With the independent case-control sampling size formula, the sampling size to detect a real odds ratio or case exposure rate with power and two-sided type I error of 5% risk was 480 including 120 HAT cases and 360 controls.

Due to the heterogeneity of the study population formed by 19 ethno-linguistic subgroups (S1 Table) and its effect on the Hardy-Weinberg equilibrium and the risk that associations results might be bias by the stratified population rather than infections due to *T. b. gambiense*, the data were firstly stratified and analysed by three major ethno-linguistic groups (Bantu, Semi-Bantu and Baka). A second analysis was performed when 10 ethno-linguistic subgroups were pooled into two groups on the basis of similarities in language spoken and their geographical proximity [45, 46]. These two groups include the Beti-Fang (Bulu, Eton, Fan, Iyassa, Kwasse, Maabi, Mvae and Ngoumba ethno-linguistic subgroups) and Wovea (Douala and Bassa ethno-linguistic subgroups) (S2 Table). To confirm results generated on the population that was stratified into major ethno-linguistic groups and two pooled ethno-linguistic subgroups, 11 ethno-linguistic subgroups derived from this stratified population including the Bamilike, Bassa, Douala, Eton, Fan, Iyassa, Kwasse, Maabi, Mvae, Mundani and Baka (S2 Table) were further separately analysed with the fisher exact test at midpoint.

Hardy-Weinberg analysis was run not only on the entire population, but also on each individual ethno-linguistic group or subgroup in order to observe the effect the population heterogeneity on HWE, the power of our study and association results. Ethno-linguistic groups or subgroups with HWE p-value deviation and less than 10 individuals or no informative data at a locus were removed for subsequent analyses.

The Cochran-Mantel-Haenszel (CMH) test implemented in PLINKv1.9 package [47] was performed with the allelic frequencies because this test can only be done with binary vars. Used as an extension of the chi-square test, the CMH test enabled to estimate the odds ratio and 95% confidence interval across the stratified populations represented here by ethno-linguistic subgroups. Using these later as covariates, it enabled to test for associations between alleles and the probability to be infected by *T. b. gambiense* and develop HAT within each ethno-linguistic subgroup. However, the CMH2 test, also implemented in PLINKv1.9 package, was used to determine if there were significant differences in the allele frequencies between different ethno-linguistic groups or subgroups. Data were visualised with R/Rstudio version 3.3.2 (2016-10-31). Results of multiple tests were adjusted by the Bonferroni correction which assumes that each of the statistical tests is independent. The significance of genotype and allele frequency differences between cases and controls within each ethno-linguistic group or subgroup were obtained and confirmed with the Fisher exact test for 2x2 contingency table.

A meta-analysis was performed on samples from ethno-linguistic subgroups that were in HWE and that showed significant association with the Fisher exact test. This was done not only on each major ethno-linguistic group, but also on the Beti-Fang and Wovea ethno-linguistic groups and the 10 ethno-linguistics subgroups mentioned above.

## Results

### Study designed

For this study, a total of 323 individuals including 73 (22.60%) HAT cases and 250 (77.40%) controls were analysed. The 323 individuals belonged to 19 different ethno-linguistic

subgroups (S1 Table). Of these 323 individuals, 211 (65.33%) were Bantu, 84 (26.01%) semi-Bantu, 21 (6.50%) Baka and 7 (2.17%) Sudano-Sao (S1 Table). The mean age (range) was  $45.75 \pm 5$  (14–91) for HAT cases and  $37.58 \pm 5$  (9–88) for controls. No significant difference ( $t = -0.206$ ,  $P = 0.837$ ) was observed between the age of controls and HAT cases. The overall sex ratio (male/female) was 1.006 with 49.85% (161/323) of female and 50.15% (162/323) of male.

For this study, we genotyped one polymorphism in each of the seven genes (*IL1A*, *IL1RN*, *IL4RN*, *IL6*, *HP*, *HPR*, and *HLA-G*) in 73 cases and 250 controls collected from 19 ethno-linguistic subgroups from three Cameroonian HAT foci.

With LD  $r^2$  of 0.7, a disease prevalence of  $<0.01$ , the disease allele frequencies of 0.052–0.500 for 7 loci genotyped, and a sampling size of 323 individuals including 73 HAT cases and 250 controls, the power of this study was estimated at 82%.

### Genes and loci selected and genotyped

Seven loci containing 3 SNPs, 3 VNTRs and one indel were tested from 7 candidate genes. The polymorphism at each of these loci was investigated on 323 samples containing 73 HAT cases and 250 controls from three HAT foci of southern Cameroon. From 323 samples that were analysed at different loci, more than 94% were successfully genotyped at each of the 7 loci. At all loci except *HLA-G*, the reference allele was at higher frequency than the alternate allele. The genotypes 1A1A (allele 1–4: 410 bp), 1A4A (allele 4–3: 335 bp) and 1A3A (allele 3–5: 500 bp) were identified for *IL1RN* gene while the genotypes 2A2A (allele 2–2: 240 bp) and 5A5A (allele 5–6: 595 bp) or their heterozygote genotypes combinations were absent in our studied population.

For the 318 samples that were successfully genotyped at the SNPs rs1800794 of *IL1A*, 97 (30.5%), 179 (56.3%) and 42 (13.2%) were respectively heterozygote, homozygote wild-type and homozygote mutant. At SNP rs1554606 of *IL6*, 98.8% (319/323) of samples were successfully genotyped: 41.4% (132/319) were heterozygote while 49.2% (157/319) and 9.4% (30/319) were homozygote wild-type and mutant respectively. At SNP rs1697370 of *HPR*, 99.4% (321/323) of samples were genotyped: 33.3% (107/321) were heterozygote whereas 57.0% (183/321) and 9.7% (31/321) were homozygote wild-type and mutant respectively.

Regarding the VNTR, 322 (99.7%: 322 /323) samples were successfully genotyped for the genes *HP* and *IL4RN*. For *HP*, 180 (55.9%: 180/322) samples were heterozygote while 82 (25.5%: 82/322) and 60 (18.6%: 60/322) were homozygote wild-type and mutant respectively. For *IL4RN* at locus rs79071878, 176 (54.7%: 176/322), 77 (23.9%: 77/322) and 69 (21.4%: 69/322) were heterozygote, homozygote wild-type and homozygote mutant respectively. At locus rs2234663 of *IL1RN*, 99.1% (320/323) of samples were genotyped: 89.1% (285/320) were homozygote wild-type and the remaining was heterozygote.

For the indel *HLA-G* at locus rs371194629, 99.1% (320/323) of samples were genotyped and 55.9% (179/323), 18.1% (58/320) and 25.9% (83/320) were heterozygote, homozygote wild-type and homozygote mutant respectively.

### Association study performed on the whole population

Allele and genotype frequencies of all cases were compared with those of all controls at all loci using chi-squared tests. No significant difference was observed for the 14 bp indel located at rs371194629 of *HLA-G*, the SNPs rs1554606 and rs1697370 of *IL6* and *HPR* respectively and the VNTRs of *IL4RN* and *HP* (Table 3). However, a significant increased risk to be infected by *T. b. gambiense* and develop HAT was observed with the TT genotype in *IL1A* gene with an OR of 2.938 (CI<sub>95</sub> [1.56–3.89]) and a *P* value of 0.0010. In addition, the genotype 1A/3A located at locus rs2234663 of *IL1RN* VNTR with an OR of 2.71 (CI<sub>95</sub> [0.97–7.58]) and a *P*

Table 3. Genotype frequency distribution for gene polymorphisms within the global Cameroon population.

Loci	Genotypes	Case (%)	Control (%)	P-value	OR (95%CI)	BONF	X <sup>2</sup>	HWE
INDEL	<b>HLA-G: rs371194629 (Case = 72, Control = 248)</b>							
	Ins/Ins	11 (15.28)	47 (18.88)	-	-	-	-	0.0749
	Del/Ins	41 (56.94)	138 (55.42)	0.360	1.27 (0.31–3.98)	1	0.837	
	Del/Del	20 (27.78)	63 (25.30)	0.758	1.36 (0.39–4.98)	1	0.554	
VNTR	<b>IL4RN: rs79071878 (Case = 73, Control = 249)</b>							
	R1/R1	12(16.44)	65 (26.12)	-	-	-	-	0.3739
	R1/R2	44 (60.27)	132 (53.01)	0.661	1.81 (0.83–4.05)	1	0.193	
	R2/R2	17 (23.29)	52 (20.88)	0.234	1.77 (0.70–4.72)	1	2.903	
	<b>HP (Case = 72, Control = 250)</b>							
	Hp1/1	14(19.45)	68 (27.20)	-	-	-	-	0.0965
	Hp1/2	42(58.33)	138 (55.20)	0.377	1.48 (0.59–2.45)	1	0.779	
	Hp2/2	16(22.22)	44 (17.60)	0.357	1.77 (0.88–4.70)	1	2.059	
	<b>IL1RN: rs2234663 (Case = 72, Control = 248)</b>							
	1A/1A	60(84.72)	209(84.27)	-	-	-	-	1
	1A/4A	5(6.95)	30(12.10)	0.217	0.58 (0.21–1.56)	1	1.521	
	1A/3A	7(8.33)	9(3.63)	0.0012	2.71 (0.97–4.08)	0.0084	21.727	
	SNP	<b>IL1A: rs1800794 (Case = 72, Control = 246)</b>						
CC		31 (43.06)	148 (60.16)	-	-	-	-	0.007
TC		25 (34.72)	72 (29.27)	0.390	1.66 (1.15–4.92)	1	1.455	
TT		16 (22.22)	26 (10.57)	0.0010	2.94(1.56–3.89)	0.0072	9.166	
<b>IL6: rs1554606 (Case = 73, Control = 248)</b>								
TT		37 (50.68)	120 (48.39)	-	-	-	-	0.6489
TG		26 (35.62)	108 (43.55)	0.734	0.66 (0.59–1.40)	1	0.115	
GG		10 (13.70)	20 (8.06)	0.152	1.21 (0.79–3.37)	1	3.766	
<b>HPR: rs1697370 (Case = 73, Control = 248)</b>								
TT		41 (56.16)	142 (57.26)	-	-	-	-	0.1322
TC		22 (30.14)	85 (34.27)	0.684	0.94 (0.63–2.04)	1	0.166	
CC		10 (13.70)	21 (8.47)	0.387	1.17 (0.85–3.53)	1	1.899	

(%): genotype frequencies; P-value: Nominal P unadjusted asymptotic probability value; OR: odds ratio; X<sup>2</sup>: Chi-square probability value; CI: Confidence Interval at 95%; N: sample size (cases/controls); SNP: Single Nucleotide Polymorphism; VNTR: Variable Number Tandem Repeats; INDEL: Insertion and Deletion; HP: Haptoglobin; HPR: Haptoglobin related protein; IL: Interleukins; HLA: Human Leukocyte Antigen

<https://doi.org/10.1371/journal.pntd.0007283.t003>

value of 0.0012 was also associated with an increased risk of getting *T. b. gambiense* infections and develop HAT. However, the frequencies of this genotype were low in both cases (8.3%) and controls (3.6%) (Table 3) and this observation should be considered provisional until replicated in larger studies because only 7 cases and 9 controls were enrolled in the analyses. The observed differences in the allelic frequencies distribution (S3 Table) and their corresponding *p* values for the 7 loci within *IL1A*, *IL6*, *HP*, *HPR*, *IL1RN*, *IL4RN* and *HLA-G* were deduced from genotypes data contained in Table 3. Although *IL1A* seems to be associated with an increased risk of getting *T. b. gambiense* infections and develop HAT, the allele frequencies were not in Hardy-Weinberg equilibrium (HWE) (0.007). However, when the population was stratified into ethno-linguistic subgroups or major ethno-linguistic groups, the allele frequencies were in HWE for most loci genotyped as shown in S4 Table. These results indicate that the heterogeneous nature of the studied population, formed by several ethno-linguistic subgroups, has an impact on the HWE. Due to these variations and the deviation of HWE in the entire population, additional analyses were performed with the Cochran-Mantel-Haentszel test

(CMH) that takes into account the population stratification. For these analyses, the population was stratified on the basis of ethno-linguistic groups and subgroups.

### Association study considering the stratified populations

After stratification of our study population into three major ethno-linguistic groups (Bantu, Semi-Bantu and Baka), the observed allelic frequencies were all in Hardy–Weinberg equilibrium within each ethno-linguistic group; suggesting random genetic exchange within each of the major ethno-linguistic groups. Data of S4 Table shows detailed results of HWE values when the population was structured into ethno-linguistic groups and in pooled ethno-linguistic subgroups.

The Cochran-Mantel-Haenszel test (CMH) was used to test the associations between the allele frequencies and the probability to be infected by *T. b. gambiense* and develop HAT. This test estimates an odds ratio and 95% confidence interval across the population using ethno-linguistic subgroups as covariant. Data of CMH test reported in Table 4 considered only 305 individuals (69 HAT cases and 236 controls) of three major ethno-linguistic groups. The null hypothesis of the Cochran-Mantel-Haenszel (CMH) test is that allele frequencies are the same in cases and controls and do not differ between populations. With the CMH test, the minor allele T of rs1800794 in *IL1A* which is located in the promoter region was significantly associated (unadjusted P = 0.0012,  $X^2 = 30.01$ , adjusted P = 0.009) with an increased risk to be infected by *T. b. gambiense* and develop this infection (Table 4). Its OR of 2.066 (CI<sub>95</sub> [1.33–3.20]) and P value of 0.009 indicate higher frequencies in cases compared to controls.

The null hypothesis of the CMH2 test is that allele frequencies are the same in each population. The CMH2 test indicated that there was no significant difference in allele frequencies between populations (P = 0.368). The Bantu major ethno-linguistic subgroups were pooled (Beti-Fang: Bulu, Eton, Fan, Iyassa, Kwasse, Maabi, Mvae, and Nguumba; and Wovea: Bassa and Douala ethno-linguistic subgroups) into two groups (S4 Table) in order to trace and spot which of these subgroups was at the centre of this effect. With CMH test, the minor allele T of

Table 4. Cochran-Mantel-Haenszel (CMH) association analysis results within the Cameroonian population.

Loci	Gene	rsid	Alleles	MAF	P-value	BONF	CHISQ	<sup>a</sup> P-value	<sup>a</sup> BONF	<sup>a</sup> CHISQ	OR (95%CI)
SNP	<i>IL1A</i>	rs1800794	*T **C	0.285	0.0012	0.009	10.45	0.368	1	30.01	2.066(1.33–3.20)
	<i>IL6</i>	rs1554606	*G **T	0.290	0.6117	1	0.258	0.654	1	17.95	0.892(0.57–1.39)
	<i>HPR</i>	rs1697370	*C **T	0.262	0.1498	1	2.074	0.016	0.115	37.08	1.395(0.88–2.20)
VNTR	<i>HP</i>	-	*Hp2 **Hp1	0.483	0.0968	0.677	2.758	0.803	1	15.38	1.411(0.94–2.12)
	<i>ILARN</i>	rs79071878	*2R **1R	0.487	0.2673	1	1.231	0.770	1	15.99	1.26(0.84–1.89)
	<i>ILIRN</i>	rs2234663	*3A *4A **1A	0.051 0.0487	0.0859 0.4976	1 1	2.95 0.46	0.693 0.260	1 1	7.3416 24.72	2.651(0.86–8.15) 0.658(0.20–2.14)
INDEL	<i>HLA-G</i>	rs371194629	*Del **Ins	0.466	0.9599	1	0.003	0.361	1	22.69	1.011(0.67–1.52)

P-value: Nominal CMH P unadjusted asymptotic probability value

\*: minor allele

\*\*: major allele; CHISQ: Chi-square probability value; OR: odds ratio; BONF: Bonferroni adjusted asymptotic p value, MAF Minor allele frequency

<sup>a</sup>: Cochran-Mantel-Haenszel for homogeneity of association across clusters (using the cmh2 test in plink), rsid: reference SNP identification code: SNP: Single Nucleotide Polymorphism; VNTR: Variable Number Tandem Repeats; INDEL: Insertion and Deletion.

<https://doi.org/10.1371/journal.pntd.0007283.t004>

rs1800794 in *IL1A* remains significantly (unadjusted  $P = 0.0005$ ,  $X^2 = 11.99$ , adjusted  $P = 0.004$ ) associated with an increased risk of getting *T. b. gambiense* and develop this infection in the Bantu major ethno-linguistic group. Its OR of 2.32 (CI<sub>95</sub> [1.44–3.37]) and a  $P$  value of 0.0005 (S2 Table) indicates higher frequencies of the allele T in cases compared to controls. After pooling some ethno-linguistic subgroups (S2 Table), the minor allele T of rs1800794 in *IL1A* with an OR of 2.40 (CI<sub>95</sub> [1.41–4.10]) and an adjusted  $P$  value of 0.009 remains significantly associated with an increased risk of getting *T. b. gambiense* infections and develop HAT within the Beti-Fang ethno-linguistic groups.

The HP2 minor allele of *HP* seems to be also significantly (unadjusted  $P = 0.0015$ ,  $X^2 = 5.90$ , adjusted  $P = 0.011$ ; OR: 3.68 (CI<sub>95</sub> [1.23–8.33])) associated with an increasing risk of getting *T. b. gambiense* infections and develop HAT within the Bassa and Douala ethno-linguistic subgroups (S2 Table). For the other genes, no significant difference was observed in the association studies as reported on whole population (Table 4).

To confirm results obtained on the stratified populations and see the impact of heterogeneous population or different ethno-linguistic groups and subgroups on the association between gene polymorphism and the risk to be infected by *T. b. gambiense* and develop HAT, meta analyses were performed on the basis of the three major ethno-linguistic groups and 11 ethno-linguistic subgroups.

### Meta analyses on the basis of ethno-linguistic groups

Of the 323 individuals belonging to the 19 different ethno-linguistic subgroups used in this study (S1 Table), 75 of them were excluded due to small population sample size (i.e less than 10 individuals), small HWE  $P$  values (S5 Table) and or loci that were not informative (low genotypes and allelic frequencies). For subsequent analyses, 249 individuals belonging to the three major ethno-linguistic groups and 11 ethno-linguistic subgroups (Bamilike, Bassa, Douala, Eton, Fan, Iyassa, Kwasse, Maabi, Mvae, Mundani and Baka) were considered for association analysis (S5 Table). The observed allelic frequencies were all in Hardy–Weinberg equilibrium within the ethno-linguistic subgroups; suggesting random genetic exchange within these ethno-linguistic groups and subgroups.

Results of meta-analysis confirmed the significant ( $P = 0.0017$ , OR = 2.305) association previously reported for SNP rs1800794 of *IL1A* (Table 5). Its OR of 2.305 (CI<sub>95</sub> [1.29–3.25]) and a

**Table 5. Meta-analysis of the Fisher exact test results within the 11 ethno-linguistic groups clusters.**

CHR	GENE	BP (GRch37)	rsid	Alleles	N	P	P(R)	OR	Q	I
2	<i>IL1A</i>	113543273	rs1800794	*T **C	8	0.0017	0.0017	2.305	0.9176	0.00
5	<i>ILARN</i>	132680584	rs79071878	*2R **1R	9	0.1299	0.1299	0.6923	0.9781	0.00
6	<i>HLA-G</i>	29830805	rs371194629	*Del **Ins	9	0.7525	0.7525	1.0793	0.8060	0.00
7	<i>IL6</i>	154426970	rs1554606	*G **T	8	0.9748	0.9361	1.0082	0.2183	26.37
16	<i>HPR</i>	35339932	rs1697370	*C **T	8	0.6692	0.6692	0.8867	0.8810	0.00
16	<i>HP</i>	-	-	*Hp2 **HP1	10	0.09482	0.09482	0.6690	0.6859	0.00

SNP: single nucleotide polymorphism, BP (GRch37): base-pair location (Reference human genome build version); rsid: reference SNP identification code

\*: minor allele

\*\*: major allele; P: fixed-effects meta-analysis asymptotic probability value; P(R): random-effects meta-analysis asymptotic probability value; OR: fixed-effects odds ratio estimate; Q: probability value for Cochrane’s Q statistics; I: I<sup>2</sup> heterogeneity index (0–100); CHR: Chromosome; N: number of valid studies for this.

<https://doi.org/10.1371/journal.pntd.0007283.t005>



P value of 0.0017 confirms the higher frequencies of allele T in cases compared to controls. The absence of significant association at different loci of other genes was also confirmed by the meta-analysis. Results generated by meta analyses on the 11 ethno-linguistic subgroups were consistent with those of the CMH test on the populations that were stratified into three major ethno-linguistic groups and pooled ethno-linguistic subgroups.

## Discussion

In this study, 7 polymorphisms within 7 genes were investigated for their association with *T. b. gambiense* infection in southern Cameroon. The most important observation is that the minor allele (T) of *IL1A* genes could influence the infections due to *T. b. gambiense* in HAT foci of southern Cameroon. Indeed, when the 19 ethno-linguistic subgroups were grouped together and considered as one population, *IL1A* seems to be associated with an increased risk of getting *T. b. gambiense* infections and develop HAT, but the allele frequencies were not in HWE (0.007). This deviation of HWE could be due to the heterogeneity of our studied population formed by 19 ethno-linguistic subgroups with some genetic differences. This heterogeneity led to a deviation of HWE resulting probably from the Wahlund effect [48] that is caused by the variance in allele frequency among subpopulations [48–50]. Indeed, in rural areas where HAT is often found, the populations are grouped according to their ethno-linguistic subgroups with very few probabilities of marriage between people from different sub-groups. This social behaviour may lead to Wahlund effect that has an impact on HWE due to the lack of genetic exchange between populations of different ethno-linguistic subgroups. In consequence, an increase of inbreeding rate, a strong genetic drift and a decrease of the genetic diversity could be observed within and between these populations [51–53]. The small sample size of some ethno-linguistic subgroups could also increase the inbreeding effect on the high variance of allele frequencies (S4 Table). The heterogeneous structure of our population may impale a strong genetic drift that changes the gene ratio of population in a random manner. Moreover, it has been reported that the genotyping error impaled by the methods used could increase the heterozygote frequency and the observation of some mutant alleles [54–57]. These hypotheses are strengthened by the differences observed for the values of HWE within and between different ethno-linguistic subgroups (S4 Table). All these factors may induce biases in the association studies and consequently, a reduction of the power of this study.

Following stratified (Table 4) and meta analyses (Table 5), the minor allele T of rs1800794 SNP of *IL1A* gene was significantly associated with an increased risk of getting *T. b. gambiense* infections and develop HAT in major ethno-linguistic groups and subgroups of southern Cameroon. Our observations of *IL1A* were not consistent with results for the same SNP in a family based linkage study in DRC [6] which showed no significant differences in transmission rates of the C and T alleles to affected children ( $p = 0.56$ ). The discrepancy between the two sets of results could be due to the study design, the small sample size (73 cases belonging to several ethno-linguistic subgroups) and the genetic differences between the DRC and Cameroon populations. Our study was a classical case-control study involving controls and HAT patients from different families while in DRC, a family base study was performed with HAT cases and controls from the same family [6]. Indeed, *IL1* participates in macrophage activation during early *T. b. gambiense* infection in mice [15] and plays a key role in the recruitment of leukocytes into the CNS through the blood-brain barrier during CNS infection [16, 17]. Due to the fact that HAT cases and controls were matched according to sex, village and their activities, they are likely subjected to equal levels of tsetse bites. Assuming their equal exposition to tsetse bites, we can speculate that the *IL1a* variant might render some individuals refractory to *T. b. gambiense* infections. This hypothesis may increase the likelihood that some infected HAT

cases might self-cure as reported in West Africa [5]. It may also increase the likelihood that some seropositive individuals could be parasitologically negative. Such individuals could be carriers of low *T. b. gambiense* load that are below the detection threshold of parasitological techniques commonly used during medical surveys. Although well confirmed and phenotyped seropositive individuals were not included in this study, it is important to mention that an association has been reported between gene polymorphism and the progression of sleeping sickness from latent infection to active disease [13].

The 1A/3A genotype (P = 0.0012 and OR of 2.71) of *IL1RN* rs2234663 VNTR was associated with an increasing risk of getting *T. b. gambiense* infections and develop HAT. Although the allelic frequency in cases and controls is low, the results obtained indicate an increasing risk to be infected and develop *T. b. gambiense* infection in HAT foci of southern Cameroon. *IL1RN* is located in the chromosome 2q14 [58, 59] with an 86 bp VNTR polymorphism in its second intron [60, 61]. *IL1RN* polymorphisms are also closely associated with the regulation of *IL1B* activity that enhance and stimulate parasite interaction and neutralisation via the complement pathways [62, 63]. It has been proposed that individuals with the *IL1RN* 2A/2A genotype have elevated levels of circulating *IL1B*. The increased *IL1B* levels result in a prolonged inflammatory response and increase the risk of pyloric gastric disease [36]. This elevated level of *IL1B* also enhances immune-modulating and stimulating effects of the *IL1* family gene on some immune response components [64]. *IL1* family genes have been shown to enhance immune-modulating and stimulating effects on the TLF components and pro/inflammatory immune response activities during HAT infection [29, 30]. Although this is the first observation of an association between *IL1RN* and *T. b. gambiense* infections, associations have been already reported between *IL1RN* variants and other diseases such as keratoconus in a Korean population [65], *H. pylori* gastric [36] and periodontitis disease [64] in the Brazilian population, and hepatitis and primary biliary cirrhosis in the Chinese population [66].

Our findings suggest an association between *IL1* family genes (*IL1RN* and *IL1A*) and the risk of getting *T. b. gambiense* infections and develop HAT in southern Cameroon. Although the mechanism leading to this infection is not well understood, the hypothesis is that the blood-brain barrier permeability is modified due to the presence, in blood and/or in CNS, of inflammatory mediators such as *IL1*, *IL6* and *TNFA* [67]. It has been shown *in vitro* that *T. b. gambiense* induces the synthesis of inflammatory and pro-inflammatory cytokines like *IL6* and *IL1* from the bone marrow endothelial cells [68]. The involvement of these cytokines, whose level increases in the cerebrospinal fluid (CSF) during *T. b. gambiense* infection and decreases after treatment, has been confirmed [69]. *IL1A* and *IL1RN* are important immunologic regulators that compete with other *IL1* family members for the *IL1* receptor and act as negative regulators with anti-inflammatory effects [70, 71] and also in differential modulation of *IL1* activity [72]. These variants are in linkage disequilibrium with other unidentified and identified variants in the *IL1* gene family and it remains to be determined which are the functional polymorphism(s).

The minor allele HP2 with a P value of 0.0015 and an OR of 3.68 is associated with an increased risk of getting *T. b. gambiense* infections and develop HAT in Bassa and Douala ethno-linguistic subgroups (S2 Table). These results are in line with those reported in South American where an increasing risk effect of Hp2/2 genotype and HP2 allele was observed for *T. cruzi* infections [73, 74]. In gastric cancer, similar results were observed for Hp2/2 genotype [75]. However, Hp2/2 genotype and Hp2 allele were suggested to be protective against infections due to *Plasmodium falciparum* during severe malaria [34]. These results indicate the involvement of Hp2/2 and HP2 in many infectious diseases. The geographical differences in the allele frequency of Hp2 could explain its association and involvement with different susceptibility to infectious diseases [76]. Indeed, *HP* is involved in scavenging haem from lysed

red blood cells and trypanosomes cause extensive lysis of red cells, leading to a decline in circulating haem. A decline in the expression of the haptoglobin receptor on macrophages indicates earliest detectable sign of infection of mice with *T. congolense*. This expression of haptoglobin receptor (*Cd163*) declined tenfold by day 3 post infection before there was a detectable parasitaemia [77].

Our results showing no significant association for SNPs rs1554606 of *IL6* and 14 bp indel at rs371194629 of *HLA-G* do not corroborate results reported in DRC and Guinea [6, 7, 13]. For the same genes, but at different loci, Courtin *et al.* [7] showed that the T allele at rs2069849 of *IL6* seems to significantly (Bonferroni corrected  $P = 0.04$ ) decrease the risk of getting *T. b. gambiense* infections in the DRC while Kabore *et al.* [13] reported an association between the allele A at rs1818879 locus of *IL6* with the low risk of HAT progression in Guinea. The discrepancies between these results could be due to insufficient linkage between our SNPs and those (rs2069849 and rs1818879) genotyped by other authors, the genetic differences between the DRC, Guinea and the Cameroon populations, the genotyping method and the sample size. The study designs also differed because we used a case control approach while Courtin *et al.* [6] used a family-based design and our small sample size has impact on the power of this study.

The results discussed above for *IL6*, *IL4* and *HLA-G* should be considered with caution because of the heterogeneity of the studied population that induced some smaller sample sizes and limited the power of this study. Despite the efforts undertaken through several large-scale field surveys conducted in three HAT foci, we were only able to collect a relatively small number of 73 HAT cases. As already indicated in our previous publication [11], our power calculations indicated that effects of the sizes could be detected with our relatively small number of samples. However, larger cohorts of well phenotyped cases and controls may be required to confirm our observations. Although the present data is only suggestive of an association, the finding of suggestive associations in multiple populations may increase the probability that these are genuine associations with sleeping sickness.

## Conclusion

This study revealed that one SNP (rs1800794) of *IL1A* and one VNTR (rs2234663) of *IL1RN* were associated with an increased risk to be infected by *T. b. gambiense* and develop HAT in inhabitants of sleeping sickness foci of southern Cameroon. The minor allele (T) of SNP rs1800794 of *IL1A* gene and the genotype 1A3A of *IL1RN* rs2234663 VNTR seem to increase the risk of getting *T. b. gambiense* infections and develop HAT in southern Cameroon. Results of this study show that the association between host genetic determinants or gene polymorphisms and the risk to be infected by *T. b. gambiense* and develop HAT may vary with the heterogeneity of the studied populations.

## Supporting information

**S1 Table. Ethno-linguistic composition of HAT cases and controls of the studied population.**

(DOCX)

**S2 Table. The observed allelic frequencies differences distribution within Cameroon population.**

(DOCX)

**S3 Table. Cochran-Mantel-Haenszel (CMH) association analysis results within the major ethnic groups and pooled ethno-linguistic subgroups within the population.**

(DOCX)

**S4 Table. Variations of HWE values according to loci and ethno-linguistic groups and subgroups.**

(DOCX)

**S5 Table. Fisher analysis results within three major ethnic groups and 10 ethno-linguistic subgroups.**

(DOCX)

## Acknowledgments

We are grateful for the support and participation of the study populations included in this study as well as the technical team of the MINSANTE, Divisional Center for Diseases, PNLTHA. We also thank the anonymous reviewers for helpful comments on the manuscript.

## Author Contributions

**Conceptualization:** Elvis Ofon, Enock Matovu, Gustave Simo.

**Data curation:** Harry Noyes.

**Formal analysis:** Elvis Ofon, Harry Noyes, Pythagore Fogue.

**Funding acquisition:** Mathurin Koffi, Christiane Hertz-Fowler, Annette MacLeod, Enock Matovu, Gustave Simo.

**Investigation:** Elvis Ofon, Vincent Ebo'o Eyanga.

**Methodology:** Elvis Ofon, Vincent Ebo'o Eyanga, Flobert Njiokou, Mathurin Koffi, Pythagore Fogue, Gustave Simo.

**Project administration:** Enock Matovu, Gustave Simo.

**Software:** Harry Noyes.

**Supervision:** Vincent Ebo'o Eyanga, Flobert Njiokou, Gustave Simo.

**Writing – original draft:** Elvis Ofon, Flobert Njiokou, Gustave Simo.

**Writing – review & editing:** Harry Noyes, Christiane Hertz-Fowler, Annette MacLeod, Gustave Simo.

## References

1. Njiokou F., Laveissière C., Simo G., Nkinin S., Grébaud P., Cuny G., Herder S. (2006). Wild fauna as a probable animal reservoir for *Trypanosoma brucei gambiense* in Cameroon. *Infection, Genetics and Evolution*, 6:147–153. <https://doi.org/10.1016/j.meegid.2005.04.003> PMID: 16236560
2. Brun R., Blum J., Chappuis F., Burri C. (2010). Human African Trypanosomiasis. *Lancet*. 375: 148–159. [https://doi.org/10.1016/S0140-6736\(09\)60829-1](https://doi.org/10.1016/S0140-6736(09)60829-1) PMID: 19833383
3. World Health Organization, (2016). Human African Trypanosomiasis (HAT). Country Leadership and Collaboration on Neglected Tropical Diseases.
4. Büscher P., Bart J.-M., Boelaert M., Bucheton B., Cecchi G., Chitnis N., et al. (2018). Do Cryptic Reservoirs Threaten Gambiense-Sleeping Sickness Elimination? *Trends in Parasitology*, 34(3), 197–207. <https://doi.org/10.1016/j.pt.2017.11.008> PMID: 29396200

5. Jamonneau V., Ilboudo H., Kabore J., Kaba D., Koffi M., Solano P., Garcia A., Courtin D., Laveissiere C., Lingue K., Buscher P., Bucheton B. (2012). Untreated human infections by *Trypanosoma brucei gambiense* are not 100% fatal. *PLoS Negl. Trop. Dis.* 6, e1691.
6. Courtin D., Milet J., Jamonneau V., Yeminanga C.S., Kumeso V.K., Bilengue C.M., Betard C., Garcia A., (2007). Association between human African trypanosomiasis and the IL6 gene in a Congolese population. *Infect. Genet. Evol.* 7, 60–68. <https://doi.org/10.1016/j.meegid.2006.04.001> PMID: 16720107
7. Courtin D, Milet J., Sabbagh A., Massaro J.D., Castelli E.C., Jamonneau V., Bucheton B., Sese C., Favier B., Rouas-Freiss N., Moreau P., Donadi E.A., Garcia A., (2013). HLA-G 30 UTR-2 haplotype is associated with Human African trypanosomiasis susceptibility. *Infection, Genetics and Evolution* 17 (2013) 1–7. <https://doi.org/10.1016/j.meegid.2013.03.004> PMID: 23541412
8. Ilboudo H., Berthier D., Camara M., Camara O., Kabore J., et al. (2012). APOL1 expression is induced by *Trypanosoma brucei gambiense* infection but is not associated with differential susceptibility to sleeping sickness. *Infect Genet Evol* 12: 1519–1523. <https://doi.org/10.1016/j.meegid.2012.05.010> PMID: 22691369
9. Hardwick J.R., Menard A., Sironi M., Milet J., Garcia A., C Sese C., Yang F., Fu B., David Courtin D., Hollox E.J. (2014). Haptoglobin (HP) and Haptoglobin-related protein (HPR) copy number variation, natural selection, and trypanosomiasis. *Hum. Genet.* 133:69–83. <https://doi.org/10.1007/s00439-013-1352-x> PMID: 24005574
10. Cooper A, Ilboudo H., Alibu V. P., Ravel S., Enyaru J., Weir W., et al. (2017). APOL1 renal risk variants have contrasting resistance and susceptibility associations with African trypanosomiasis. *eLife*, 6, 56. <http://doi.org/10.7554/eLife.25461>
11. Ofon E, Noyes H., Mulindwa J., Ilboudo H., Simuunza M., Ebo'o V., et al. (2017). A polymorphism in the haptoglobin, haptoglobin related protein locus is associated with risk of human sleeping sickness within Cameroonian populations. *PLoS Negl Trop Dis* 11(10):e0005979. <https://doi.org/10.1371/journal.pntd.0005979> PMID: 29077717
12. Ahouty B, Koffi M., Ilboudo H., Simo G., Matovu E., Mulindwa J., et al. (2017). Candidate genes-based investigation of susceptibility to Human African Trypanosomiasis in Cote d'Ivoire. *PLoS Negl Trop Dis* 11(10): e0005992. <https://doi.org/10.1371/journal.pntd.0005992> PMID: 29059176
13. Kaboré J.W., Ilboudo H., Noyes H., Camara O., Kaboré J., Camara M., et al. (2017). Candidate gene polymorphisms study between human African trypanosomiasis clinical phenotypes in Guinea. *PLoS Negl Trop Dis* 11(8): e0005833. <https://doi.org/10.1371/journal.pntd.0005833> PMID: 28827791
14. Gineau L, Courtin D, Camara M, Ilboudo H, Jamonneau V, Dias FC, Tokplonou L, Milet J, Mendonça PB, Castelli EC, Camara O, Camara M, Favier B, Rouas-Freiss N, Moreau P, Donadi EA, Bucheton B, Sabbagh A, Garcia A. (2016). Human Leukocyte Antigen-G: A Promising Prognostic Marker of Disease Progression to Improve the Control of Human African Trypanosomiasis. *Clin Infect Dis.* 2016 Nov 1; 63 (9):1189–1197. <https://doi.org/10.1093/cid/ciw505> PMID: 27470243
15. Sternberg J. M. (2004). Human African trypanosomiasis: clinical presentation and immune response. Review Article. *Parasite Immunology*, 2004, 26, 469–476 <https://doi.org/10.1111/j.0141-9838.2004.00731.x> PMID: 15771682
16. Borges L.F. (1992). Infections in neurologic surgery. Host defenses. *Neurosurg. Clin. N. Am.* 3, 275–278 PMID: 1633459
17. Ching S, He L., Lai W., Quan N. (2005). IL-1 type I receptor plays a key role in mediating the recruitment of leukocytes into the central nervous system. *Brain Behav Immun* 19:127–137. <https://doi.org/10.1016/j.bbi.2004.06.001> PMID: 15664785
18. Penchenier L., Grébaut P., Ebo'o Eyenga V., Bodo J.M., Njiokou F., Binzouli J.J., Simarro P., Soula G., Herder S. (1999). Le foyer de la trypanosomiase humaine africaine de Campo (Cameroun) en 1998. Aspects épidémiologiques, état de l'endémie et comparaison des CATT 1.3 et CATT Latex dans le dépistage de masse. *Bull. Soc. Path. Ex.* 92, 185–190.
19. Simo G, Mbida J.A., Eyenga V.E., Asonganyi T., Njiokou F., Grébaut P. (2014). Challenges towards the elimination of Human African trypanosomiasis in the sleeping sickness focus of Campo in southern Cameroon. *Parasites Vectors* 7, 374. <https://doi.org/10.1186/1756-3305-7-374> PMID: 25129168
20. Assonganyi T, Bedifeh B.A., Ade S.S., Ngu J.L. (1994). An evaluation of the card agglutination test for trypanosomiasis (CATT) reagent in the Fontem sleeping Sickness focus, Cameroon. *Afric. J. Med. Med. Sci.* 23(1)[39–46].
21. Mélachio T.T., Simo G., Ravel S., De Meeûs T., Causse S., Solano P., Lutumba P., Asonganyi T., Njiokou F. (2011). Population genetics of *Glossina palpalis palpalis* from central African sleeping sickness foci. *Parasites Vectors* 4, 140. <https://doi.org/10.1186/1756-3305-4-140> PMID: 21767402
22. Simo G, Asonganyi T., Nkinin S.W., Njiokou F., Herder S. (2006). High prevalence of *Trypanosoma brucei gambiense* group 1 in pigs from the Fontem sleeping sickness focus in Cameroon. *Veterinary Parasitology*, 139: 57±66. <https://doi.org/10.1016/j.vetpar.2006.02.026> PMID: 16567049



23. Kouega J.-P. (2007). The Language Situation in Cameroon. *Current Issues in Language Planning*, 8 (1), 3–93. <http://doi.org/10.2167/cilp110.0>
24. Magnus E, Vervoort T., Van Meirvenne N. (1978). A card-agglutination test with stained trypanosomes (C.A.T.T.) for the serological diagnosis of *T. b. gambiense* trypanosomiasis. *Ann. Soc. Belg. Méd. Trop.* 58: 169–176. PMID: 747425.
25. WHO/World Health Organization, 1986. African trypanosomiasis: epidemiology and fight against. *Technical Rapport Series*, 739, 32pages.
26. Buscher P, Mumba N.D., Kabore J., Lejon V., Robays J., Jamonneau V., Bebronne N., Van der Veken W., Bieler S. (2009). Improved models of minianion exchange centrifugation technique (mAECT) and modified single centrifugation (MSC) for sleeping sickness diagnosis and staging. *PLoS Negl. Trop. Dis.* 3, e471. <https://doi.org/10.1371/journal.pntd.0000471> PMID: 19936296
27. Jamonneau V, Bucheton B., Kabore J., Ilboudo H., Camara O., et al., (2010). Revisiting the immune trypanolysis test to optimise epidemiological surveillance and control of sleeping sickness in West Africa. *PLoS Negl Trop Dis* 4: e917. <https://doi.org/10.1371/journal.pntd.0000917> PMID: 21200417
28. Van Meirvenne N, Magnus E, Buscher P. 1995. Evaluation of variant specific trypanolysis tests for serodiagnosis of human infections with *Trypanosoma brucei gambiense*. *Acta Tropica*, 60(3), 189–199. PMID: 8907397
29. Pays E, Vanhollenbeke B, Vanhamme L, Paturiaux-Hanocq F, Nolan DP, et al. (2006). The trypanolytic factor of human serum. *Nat. Rev. Microbiol.* 4: 477–486. <https://doi.org/10.1038/nrmicro1428> PMID: 16710327
30. Checchi F, Filipe J.A., Barrett M.P. & Chandramohan D. (2008). The natural progression of Gambiense sleeping sickness: what is the evidence? *PLoS Negl Trop Dis*; 2: e303. <https://doi.org/10.1371/journal.pntd.0000303> PMID: 19104656
31. Bucheton B, MacLeod A., Jamonneau V. (2011). Human host determinants influencing the outcome of *Trypanosoma brucei gambiense* infections. *Parasite Immunol* 33: 438±447. <https://doi.org/10.1111/j.1365-3024.2011.01287.x> PMID: 21385185
32. Sternberg JM & Maclean L. (2010). A spectrum of disease in Human African trypanosomiasis: the host and parasite genetics of virulence. *Parasitology*; 137: 2007–2015. <https://doi.org/10.1017/S0031182010000946> PMID: 20663245
33. Ilboudo H, Bras-Goncalves R., Camara M., Flori L., Camara O., Sakande H., Leno M., Petitdidier E., Jamonneau V., Bucheton B. (2014). Unravelling Human Trypanotolerance: IL8 is Associated with Infection Control whereas IL10 and TNF $\alpha$  Are Associated with Subsequent Disease Development. *PLoS-Pathog* 10(11): e1004469. <https://doi.org/10.1371/journal.ppat.1004469> PMID: 25375156
34. Minang J.T., Gyan B.A., Anchang J.K., Troye-Blomberg M., Perlmann H., Achidi E.A. (2004). Haptoglobin phenotypes and malaria infection in pregnant women at delivery in western Cameroon. *Acta Trop.* 90, 107–114. PMID: 14739029
35. Courtin D, Berthier D, Thevenon S., Dayo G-K., Garcia A., Bucheton B. (2008). Host genetics in African trypanosomiasis. *Infect. Gen. Evol.* 8:229–238.
36. Santos J.C., Ladeira M.S.P., Pedrazzoli J. Jr. & Ribeiro M.L. (2012). Relationship of *IL-1* and *TNF- $\alpha$*  polymorphisms with *Helicobacter pylori* in gastric diseases in a Brazilian population. *Braz. J. Med. Biol. Res.*, 45(9) 811–817. <https://doi.org/10.1590/S0100-879X2012007500099> PMID: 22714811
37. Tripathi A.K., Shukla S., Tripathi J.K., Saket R.D., Kol S., Mishra P., Chauhan U.K., and Indurkar M. (2015). Association of Genetic Polymorphism of Inflammatory Genes (IL-1 $\beta$  and IL-4) with Diabetes Type 2. *Enliven: J Genet Mol Cell Biol* 1(1): 004.
38. Koch W, Latz W., Eichinger M., Roguin A., Levy A.P., Schomig A., Kastrati A. (2002). Genotyping of the common haptoglobin Hp 1/2 polymorphism based on PCR. *Clin Chem* 48:1377–1382 PMID: 12194911
39. Koressaar T. and Remm M. (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics*; 23(10):1289–91. <https://doi.org/10.1093/bioinformatics/btm091> PMID: 17379693
40. Untergasser A, Cutcutache I., Koressaar T., Ye J., Faircloth B.C., Remm M. and Rozen S.G. (2012). Primer3—new capabilities and interfaces. *Nucleic Acids Res.* 2012 Aug 1; 40(15):e115. <https://doi.org/10.1093/nar/gks596> PMID: 22730293
41. Castelli E.C., Mendes-Junior C.T., Deghaide N.H., de Albuquerque R.S., Muniz Y.C., Simoes R.T., Carlosella E.D., Moreau P., Donadi E.A., (2010). The genetic structure of 3' untranslated region of the HLA-G gene: polymorphisms and haplotypes. *Genes Immun.* 11, 134–141. <https://doi.org/10.1038/gene.2009.74> PMID: 19798077
42. Menashe I., Rosenberg P. S. & Chen B. E. (2008). PGA: power calculator for case-control genetic association analyses. *BMC Genet* 9, 36. <https://doi.org/10.1186/1471-2156-9-36> PMID: 18477402

43. Franco JR, Cecchi G, Priotto G, Paone M, Diarra A, Grout L, Mattioli R, Argaw D. (2017). Monitoring the elimination of human African trypanosomiasis: Update to 2014. *PLoS Neglected Tropical Disease* 11 (5): e0005585. <https://doi.org/10.1371/journal.pntd.0005585>
44. Kasiulevicius V, Sapoka V, Filipaviciute R. (2006). Sample size calculation in epidemiological studies. *Gerontology*. 2006; 7:225–31.
45. Barrier J.C. (1983). Les groupes ethniques et les langues: ORSTOM fonds documentaire N0: 02981, MEA, H: 239–260
46. Mohamad Z.Y (1999). African peoples and Nations by country: Cameroon. *Almanac of African peoples and Nations*, Transaction Publisher, New Brunswick N.J. p37-39.
47. Purcell S, Neale B., Todd-Brown K., Thomas L., Ferreira M.A.R, Bender D., Maller J., Sklar P., de Bakker P.I.W, Daly M.J. & Sham P.C. (2007). PLINK: a toolset for whole-genome association and population-based linkage analysis. *Amer. J. Hum. Genet.* 81.
48. Wahlund S (1928). Zusammensetzung von Populationen und Korrelationserscheinungen vom Standpunkt der Vererbungslehre aus betrachtet. *Hereditas* 11: 65–106.
49. Sinnock P (1975). The Wahlund effect for the two locus model. *Am. Nat.* 109: 565–570.
50. Hartl DL and Clark AG (2007). *Principles of Population Genetics*, 4th Edn. Sunderland publisher Editorial board, MA: Sinauer Associates.
51. Freedman ML, Reich D, Penney KL et al. (2004). Assessing the impact of population stratification on genetic association studies. *Nature Genetics* 36: 388–393. <https://doi.org/10.1038/ng1333> PMID: 15052270
52. Marchini J, Cardon LR, Phillips MS et al. (2004). The effects of human population structure on large genetic association studies. *Nature Genetics* 36: 512–517. <https://doi.org/10.1038/ng1337> PMID: 15052271
53. Chikhi L, Sousa V, Luisi P, Goossens B, Beaumont MA (2010). The confounding effects of population structure, genetic diversity and the sampling scheme on the detection and quantification of population size changes. *Genetics* 186: 983–995. <https://doi.org/10.1534/genetics.110.118661> PMID: 20739713
54. Zintzaras E, Lau J. (2008). Synthesis of genetic association studies for pertinent gene disease associations requires appropriate methodological and statistical approaches. *J Clin Epidemiol.* 61:634–45. <https://doi.org/10.1016/j.jclinepi.2007.12.011> PMID: 18538260
55. Sillanpää MJ (2011). Overview of techniques to account for confounding due to population stratification and cryptic relatedness in genomic data association analyses. *Heredity* 106: 511–519. <https://doi.org/10.1038/hdy.2010.91> PMID: 20628415
56. Chen B, Cole JW and Grond-Ginsbach C (2017). Departure from Hardy Weinberg Equilibrium and Genotyping Error. *Front. Genet.* 8:167. <https://doi.org/10.3389/fgene.2017.00167> PMID: 29163635
57. Wittke-Thompson JK, Pluzhnikov A, Cox NJ (2005). Rational inferences about departures from Hardy-Weinberg equilibrium. *Am J Hum Genet.* 2005; 76: 967–86. <https://doi.org/10.1086/430507> PMID: 15834813
58. Arend WP (2002). The balance between IL-1 and IL-1Ra in disease. *Cytokine Growth Factor Rev* 13:323–340 PMID: 12220547
59. Qian N, Chen X., Han S., Qiang F., Jin G., Zhou X., Dong J., Wang X., Shen H., Hu Z. (2010). Circulating IL-1b levels, polymorphisms of IL-1B, and risk of cervical cancer in Chinese women. *J Cancer Res Clin Oncol* 136:709–716 <https://doi.org/10.1007/s00432-009-0710-5> PMID: 19904560
60. Zhang Y, Liu C., Peng H., Zhang J., Feng Q. (2012). IL1 receptor antagonist gene IL1-RN variable number of tandem repeats polymorphism and cancer risk: a literature review and metaanalysis. *PLoS One* 7:e46017 <https://doi.org/10.1371/journal.pone.0046017> PMID: 23049925
61. Tong Y, Han J., Guan X., Lu Z., Miao X., Ye J., Hou S.Y., Zhang Y., Geng Y., Li Y., Zhan F., Liu J., Cheng J.Q. (2013). Association of IL-1 receptor antagonist gene VNTR polymorphism with ischemic stroke in the Chinese Uyghur population. *Biochem Genet* 51:698–706 <https://doi.org/10.1007/s10528-013-9599-4> PMID: 23743874
62. El-Omar E.M., Oien K., Murray L.S., El-Nujumi A., Wirz A., Gillen D., William C., Fullarton G., McColl K. E. (2000). Increased prevalence of precancerous changes in relatives of gastric cancer patients: critical role of *H. pylori*. *Gastroenterology* 2000; 118: 22–30. PMID: 10611150
63. Camargo M.C., Mera R., Correa P., Peek R.M. Jr., Fontham E.T., Goodman K.J., Piazuelo M.B., Sicinski L., Zabaleta J., Schneider B.G. (2006). Interleukin-1beta and interleukin-1 receptor antagonist gene polymorphisms and gastric cancer: a meta-analysis. *Cancer Epidemiol Biomarkers Prev.* 15 (9):1674–87. <https://doi.org/10.1158/1055-9965.EPI-06-0189> PMID: 16985030
64. Trevilatto P.C., de Souza-Pardo A.P., Scarel-Caminaga R.M., de Brito R.B. Jr., Alvim-Pereira F., Alvim-Pereira C.C., Probst C.M., Garlet G.P., Sallum A.W., Line S.R.P. (2011). Association of IL1 gene

- polymorphisms with chronic periodontitis in Brazilians. *Archives of Oral Biology* 56 (2011) 54–62. <https://doi.org/10.1016/j.archoralbio.2010.09>
65. Kim S.H., J.W., Kim H.S., Joo C.K. (2008). Association of –31T>C and –511 C>T polymorphisms in the interleukin 1 beta (IL1B) promoter in Korean keratoconus patients. *Molecular Vision* 2008; 14:2109–2116 PMID: [19043479](https://pubmed.ncbi.nlm.nih.gov/19043479/)
  66. Fan L.Y., Tu X.Q., Zhu Y., Pfeiffer T., Feltens R., Stoecker W., Zhong R.Q. (2005). Genetic association of cytokines polymorphisms with autoimmune hepatitis and primary biliary cirrhosis in the Chinese. *World J Gastroenterol.* 11(18): 2768–2772 <https://doi.org/10.3748/wjg.v11.i18.2768> PMID: [15884119](https://pubmed.ncbi.nlm.nih.gov/15884119/)
  67. De Vries H.E., Blom-Roosemalen M.C., van Oosten M., de Boer A.G., van Berkel T.J., Bremier D.D. Kupier J. (1996). The influence of cytokines on the integrity of the blood-brain barrier in vitro. *J Neuroimmunol* 64: 37±43. PMID: [8598388](https://pubmed.ncbi.nlm.nih.gov/8598388/)
  68. Girard M, Giraud S., Courtioux B., Jauberteau-Marchan M.O., Bouteille B. (2005). Endothelial cell activation in the presence of African trypanosomes. *Mol Biochem Parasitol* 139: 41±49. <https://doi.org/10.1016/j.molbiopara.2004.09.008> PMID: [15610818](https://pubmed.ncbi.nlm.nih.gov/15610818/)
  69. Lejon V., Lardon J., Kenis G., Pinoges L., Legros D., Bisser S., N'Siesi X., Bosmans E., Buscher P. (2002). Interleukin (IL)-6, IL-8 and IL-10 in serum and CSF of *Trypanosoma brucei gambiense* sleeping sickness patients before and after treatment. *Trans. R. Soc. Trop. Med. Hyg.* 96, 329–333. PMID: [12174791](https://pubmed.ncbi.nlm.nih.gov/12174791/)
  70. Korthagen N.M., van Moorsel C.H.M., Kazemier K.M., Ruven H.J.T. and Grutters J.C. (2012). "IL1RN-genetic variations and risk of IPF: a meta-analysis and mRNA expression study," *Immunogenetics*, vol. 64, no. 5, pp. 371–377, 2012. <https://doi.org/10.1007/s00251-012-0604-6> PMID: [22322675](https://pubmed.ncbi.nlm.nih.gov/22322675/)
  71. Hashemi M, Naderi M., Ebrahimi M., Amininia S., Bahari G., Taheri M., Eskandari-Nasab E., and Ghavami S. (2015). Association between Interleukin-1 Receptor Antagonist (IL1RN) Variable Number of Tandem Repeats (VNTR) Polymorphism and Pulmonary Tuberculosis. *Iran J Allergy Asthma Immunol.* February 2015: 14(1):55–59. PMID: [25530139](https://pubmed.ncbi.nlm.nih.gov/25530139/)
  72. Danis V.A., Millington M., Hyland V.J., Grennan D. (1995). Cytokine production by normal human monocytes: inter-subject variation and relationship to an IL-1 receptor antagonist (IL-1Ra) gene polymorphism. *Clin Exp Immunol* 99:303–10. PMID: [7851026](https://pubmed.ncbi.nlm.nih.gov/7851026/).
  73. Calderoni D.R., Andrade S. and Grotto H.Z. (2006). "Haptoglobin phenotype appears to affect the pathogenesis of American trypanosomiasis," *Annals of Tropical Medicine and Parasitology*, vol. 100, no. 3, pp. 213–221, 2006.
  74. Kumar DM, Thota B, Shinde SV, Prasana KV, Hegde AS, Arivazhagan A, Chandramouli BA, Santosh V, Somasundaram K. (2010) Proteomic identification of haptoglobin  $\alpha 2$  as a glioblastoma serum biomarker: implication in cancer cell migration and tumor growth. *J Proteome Res.* 5; 9(11): 5557–67. <https://doi.org/10.1021/pr1001737> PMID: [20822092](https://pubmed.ncbi.nlm.nih.gov/20822092/)
  75. Jayanthi M, Habibullah CM, Ishaq M, Ali H, Babu PS, Ali MM. (1989). Distribution of haptoglobin phenotypes in oesophageal and gastric cancer. *J Med Genet.* 1989; 26: 172–3. PMID: [2709394](https://pubmed.ncbi.nlm.nih.gov/2709394/)
  76. Langlois MR and Delanghe JR. (1996). Biological and clinical significance of haptoglobin in human. *Clin Chem.* 42: 1589–600. PMID: [8855140](https://pubmed.ncbi.nlm.nih.gov/8855140/)
  77. Noyes H.A., Alimohammadian M.H., Agaba M., Brass A., Fuchs H., Gailus-Durner V., et al. (2009). Mechanisms Controlling Anaemia in *Trypanosoma congolense* Infected Mice. *PLoS ONE*, 4(4), e5170. <https://doi.org/10.1371/journal.pone.0005170> PMID: [19365556](https://pubmed.ncbi.nlm.nih.gov/19365556/)